Multigradients and the Zeros of Transcendental Functions

A. S. HOUSEHOLDER
The University of Tennessee
Knoxville, Tennessee
AND
Oak Ridge National Laboratory
Oak Ridge, Tennessee

Stewart [4] has recently obtained a quite general convergence theorem that, at one stroke, achieves several seemingly diverse objectives. First, and most directly, it extends the algorithm of Sebastião e Silva, suitably transformed, and shows it to be effective for finding the zeros, not only of polynomials, but also of transcendental functions in a circle of analyticity. Second, it generalizes the Koenig-Hadamard theorem, on which rests the justification for both the *qd* algorithm and Schroeder-type iterations [1]. Third, it generalizes the method on which the *qd* algorithm is based and provides an alternative justification. Hence it shows that the algorithm of Sebastião e Silva and the *qd* algorithm, as well as those of an infinite family of other possible ones, all rest on the same mathematical foundations.

The present purpose is to introduce a class of determinants, or rather two classes, that will be called multigradients, which are a natural generalization of bigradients, and to show that in terms of these multigradients the several algorithms have a rather simple, and uniform, representation. As bigradients are determinants formed from two polynomials or expansions and their coefficients, multigradients are formed from an arbitrary number of such. They lead to certain linear combinations that play a fundamental role in the general class of Stewart algorithms.

As indicated above, Stewart extends a certain (simple and natural) transformation of the algorithm of Sebastião e Silva rather than the algorithm itself. The original algorithm applied to a polynomial provides in sequence the zeros of this polynomial beginning with that of largest modulus, or provides polynomials of generally lower degree, each having zeros of equal modulus. One could attempt to extend this to transcendental functions that are analytic everywhere outside a sufficiently large circle, assuming the function to be given by its expansion in powers of z^{-1}. But the polynomial, say $f(z)$, if of degree n, could be replaced by $z^n f(z^{-1})$ [3], and it is the algorithm so transformed that is extended by Stewart.

Given the formal expansions

$$q_i(z) = \sum_0^\infty b_{ij} z^j, \qquad i = 0, 1, 2, \ldots, \tag{1}$$

multigradients of two types will be defined. Those of the first type are constants:

$$\delta[g_0, g_1, \ldots, g_p] = \delta \begin{bmatrix} b_{00} & b_{01} & \cdots & b_{0p} \\ b_{10} & b_{11} & \cdot & b_{1p} \\ b_{p0} & b_{p1} & \cdots & b_{pp} \end{bmatrix},$$

where the δ signifies the determinant. Those of the second type are expansions:

$$\delta[g_0(z), g_1(z), \ldots, g_p(z)] = \delta \begin{bmatrix} b_{00} & b_{01} & \cdots & b_{0,p-1} & g_0(z) \\ b_{10} & b_{11} & \cdot & b_{1,p-1} & g_1(z) \\ & & \cdot & & \\ b_{p0} & b_{p1} & \cdots & b_{p,p-1} & g_p(z) \end{bmatrix}, \tag{3}$$

differing from those of the first type only in that in the last column the $g_i(z)$ appear. By an obvious determinantal reduction (subtracting from the last column previous ones multiplied by suitable powers of z), it follows that

$$\delta[g_0(z), g_1(z), \ldots, g_p(z)] = z^p \delta[g_0, g_1, \ldots, g_p] + O(z^{p+1}). \tag{4}$$

On the other hand, expansion of the determinant on the right of (3) by elements of the last column shows it to be a linear combination with constant coefficients of the $g_i(z)$. Hence this is a linear combination of $g_0(z), \ldots, g_p(z)$ that is $O(z^p)$ and, if the multigradient in (2) is nonvanishing, any such linear combination is a constant multiple of this.

Suppose the multigradient in (2) is nonvanishing. Then the reduced linear combination (RLC) of $g_0(z), \ldots, g_p(z)$ is defined to be

$$\{g_0, g_1, \ldots, g_p\} = z^{-p} \frac{\delta[g_0(z), \ldots, g_p(z)]}{\delta[g_0, \ldots, g_p]}. \tag{5}$$

More generally, consider the matrices

$$P_\nu = \begin{bmatrix} b_{00} & b_{01} & \cdots & b_{0p} & b_{0,p+1} & \cdots & b_{0,p+\nu} \\ b_{10} & b_{11} & \cdot & b_{1p} & b_{1,p+1} & \cdot & b_{1,p+\nu} \\ b_{p0} & b_{p1} & \cdot & b_{pp} & b_{p,p+1} & \cdot & b_{p,p+\nu} \end{bmatrix}. \tag{6}$$

If P_0 is nonsingular, the set $g_0, \ldots, g_p$ will be said to be *normal*, and (5) holds. If every P_ν is of rank $\leqslant p$, the expansions are linearly dependent, and the set will be said to be *singular*, or to have *index* ∞. If this is not the case, there will be a first, say P_ν, that is of rank $p+1$, and the set will be said to have *index* ν. Hence a normal set has index 0. If the set has index ν, then there exists a unique linear combination that is equal to $z^{\nu+p} + 0(z^{\nu+p+1})$. *If the set $g_0(z), \ldots, g_p(z)$ is singular, its reduced linear combination will be defined to be*

$$\{g_0, \ldots, g_p\} = 0. \tag{5'}$$

If it has index $\nu < \infty$, it is defined to be z^{-p} times that unique linear combination such that

$$\{g_0, \ldots, g_p\} = z^\nu + 0(z^{\nu+1}). \tag{5''}$$

The coefficients of this linear combination are proportional to the determinants of certain submatrices of order p, not all vanishing, from the matrix p_ν. For a set containing only one member, that member itself will be said to be normal or to have index ν, as the case may be.

Certain elementary properties of the RLC are at once obvious. The RLC is unchanged by any permutation of the expansions in the set. It is also unchanged when to any member of the set is added any linear combination of other members and, further, it is unchanged when any member is multiplied by any nonnull constant. Somewhat less obvious are the identities of the foll wing two lemmas.

LEMMA 1. *Let the set $f_0(z), f_1(z), \ldots, f_q(z)$ be normal. Then*

$$\{f_0, f_1, \ldots, f_q, g_0, g_1, \ldots, g_p\} = \{\{f_0, \ldots, f_q, g_0\}, \ldots, \{f_0, \ldots, f_q, g_p\}\}.$$

Proof. Consider the infinite rectangular matrix of $p+q+2$ rows formed from the coefficients of the $f(z)$ and the $g(z)$. In view of the normality of the set of the $f(z)$, there exists a unit lower triangular matrix of order $p+q+2$ of the form

$$\begin{bmatrix} I & 0 \\ P & I \end{bmatrix},$$

which, when it multiplies the infinite rectangular matrix on the left, produces zeros in the first $q + 1$ elements of the last $p + 1$ rows. This does not change the rank of any submatrix formed from all the rows. Since the upper left submatrix of order $q + 1$ is nonsingular by hypothesis, the rank of such a submatrix is equal to $q + 1$ plus the rank of the submatrix from which the first $q + 1$ rows are omitted.

LEMMA 2. *If $f(z)$ is normal, then*

$$\{f, zf, \ldots, z^{\nu}f, g\} = \{f, \{f, zf, \ldots, z^{\nu-1}f, g\}\}$$

$$\{f, \{f, \{f, \ldots, \{f, g\}, \ldots\}\},$$

where f occurs $\nu + 1$ times.

Proof. The proof is made inductively, starting with the infinite matrix associated with the left member.

The algorithm of Sebastião e Silva in the form to be considered here is as follows. Let

$$f(z) = 1 + a_1 z + a_2 z^2 + \cdots \tag{7}$$

be a polynomial of degree n, and let

$$g(z) = 1 + b_1 z + b_2 z^2 + \cdots \tag{8}$$

be a polynomial of degree $n - 1$ at most, having no zero in common with $f(z)$. Let

$$g_1(z) = g(z),$$

and form $g_{\nu+1}(z)$ by eliminating the constant term between $g_\nu(z)$ and $f(z)$ and removing the factor z. The first part of the theorem states that, if $f(z)$ has zeros $r_1, r_2, \ldots$ satisfying

$$|r_1| < |r_2| \leqslant \cdots,$$

then

$$\lim_{\nu\to\infty} \frac{g_\nu(z)}{g_\nu(0)} = \frac{f(z)}{1 - z/r_1}.$$

Stewart's theorem states in part that this is still true if $f(z)$ and $g(z)$ are merely analytic in some circle about the origin containing r_1.

Moreover, let

$$g_{\nu,1}(z) = g_\nu(z),$$

and form $g_{\nu,p+1}(z)$ by eliminating the constant terms between $g_{\nu,p}(z)$ and $g_{\nu-1,p}(z)$ and dividing by z. Then, if $r_1, r_2, \ldots, r_p$ are all within the circle of analyticity, and all other zeros, if any, are strictly greater in modulus, then

$$\lim_{\nu\to\infty} \frac{g_{\nu,p}(z)}{g_{\nu,p}(0)} = \frac{f(z)}{(1 - z/r_1)\cdots(1 - z/r_p)}. \tag{9}$$

The requirement that $f(0) = g(0) = 1$ is, of course, only a convenience. However, it is also convenient, and certainly imposes no restriction, if each $g_{\nu,p}(z)$ is normalized so that $g_{\nu,p}(0) = 1$. Then the $g_{\nu,p}(z)$ are RLC's of other g's, or of f and a g, and application of Lemmas 1 and 2 shows that

$$g_{\nu,p}(z) = \{z^{\nu-1}f, \ldots, f, g, \ldots, z^{p-1}g\}. \tag{10}$$

If these functions are formed by the algorithm described above, and arranged in a doubly infinite array with ν increasing downward and p increasing to the right, it is easy to see that only functions on and below the main diagonal, i.e., with $\nu \geqslant p$, will appear. However, (10) defines also the $g_{\nu,p}(z)$ that lie above this diagonal. It is convenient further to border this table so as to obtain the following:

$$\begin{array}{cccc}
 & g & g & \cdots \\
f & \{f, g\} & \{f, g, zg\} & \cdots \\
f & \{zf, b, g\} & \{zf, f, g, zg\} & \cdots \\
\cdots & \cdots & \cdots & \cdots
\end{array}$$

This is now seen to be the table of normalized residuals in the Padé table for $g(z)/f(z)$. By this is meant the following: In general, for any ν and p, there exist polynomials $\phi_\nu(z)$ and $\psi_p(z)$, of degrees ν and p, respectively, such that

$$\psi_p(z)g(z) - \phi_\nu(z)f(z) = O(z^{\nu+p+1}). \tag{11}$$

and these are uniquely determined up to a multiplicative constant. The quotient $\phi_\nu(z)/\psi_p(z)$ is the (ν, p) entry in the Padé table for $g(z)/f(z)$, and the right-hand side of (11) is the residual. The normalized residual is obtained by dividing out the power of z, and then the constant term. It can be shown [2] that this is $\{z^\nu f, \ldots, f, g, \ldots, z^p g\}$.

If the relations of the form (11) are written for four adjacent entries in the Padé table located at vertices of a square, it can be shown that any three of the residuals satisfy a certain three-term identity, there being thus four such identities. These are among the Frobenius identities. They imply similar, but somewhat simpler, identities, each connecting three of the reduced residuals, and in these four identities only three distinct constant coefficients occur, one each in three of the identities, all three in the remaining one [2]. Of these three, two are a q and an e of the qd algorithm. Thus the algorithm described above turns out to lead directly to the qd algorithm, and conversely.

The necessary convergence proofs for the qd algorithm is ordinarily made by way of the Koenig-Hadamard theorem [1]. However, Stewart's theorem proves the convergence of algorithms of a much broader class, hence can be regarded as a generalization of the Koenig-Hadamard theorem. Consider a set of functions

$$g_1(z), g_2(z), \ldots, g_p(z),$$

in number equal to the number of zeros to be found. These are required to be analytic throughout the region of interest, to have no zero in common with $f(z)$, and to satisfy another mild restriction to be given shortly. Define the operator F by

$$Fg(z) = \{f, g\}, \tag{12}$$

(this differs unessentially from that defined by Stewart), and let

$$g_j^{(0)}(z) = g_j(z) \tag{13}$$

and

$$g_j^{(\nu)}(z) = F^\nu g(z). \tag{14}$$

Let

$$u_p^{(\nu)}(z) = \{g_1^{(\nu)}, \ldots, g_p^{(\nu)}\}. \tag{15}$$

The mild restriction referred to is that, for sufficiently large ν, the sets $g_1^{(\nu)}(z), \ldots, g_p^{(\nu)}(z)$ must all be normal. If $f(z)$ has the zeros $r_1, r_2, \ldots, r_p$ within the regions of analyticity and, if all other zeros, if any, are strictly greater in modulus, then

$$\lim_{\nu \to \infty} u_p^{(\nu)}(z) = \frac{f(z)}{(1 - z/r_1) \cdots (1 - z/r_p)}.$$

This is the substance of Stewart's theorem. More precisely, if $f(z), g_1(z), \ldots, g_p(z)$ are all analytic in the open disk of radius R, if this disk contains no other zero of $f(z)$, and if

$$|r_1| \leqslant |r_2| \leqslant \cdots \leqslant |r_p| < \rho < R,$$

then

$$u_p^{(\nu)}(z) = \frac{f(z)}{(1 - z/r_1) \cdots (1 - z/r_p)} + O(r_p^{\nu}/\rho^{\nu}). \tag{16}$$

By applying Lemmas 1 and 2 it is found that in place of (15) can be written

$$u_p^{(\nu)}(z) = \{z^{\nu-1}f, z^{\nu-2}f, \ldots, f, g_1, \ldots, g_p\}. \tag{17}$$

Hence, if

$$g_j(z) = z^{j-1}g(z),$$

these are precisely the reduced residuals for the Padé table.

In the general case when the g's are arbitrary, only one of the four identities is valid for relating adjacent entries in the table of the u's. However, by applying the Sylvester identity [1] to the multigradient

$$\delta\{z^{\nu-1}f(z), \ldots, f(z), g_1(z), \ldots, g_p(z)\},$$

it is found that

$$r_p^{(\nu)} z u_p^{(\nu)}(z) = u_p^{(\nu-1)}(z) - u_{p-1}^{(\nu)}(z), \tag{18}$$

where

$$r_p^{(\nu)} = \frac{\delta\{z^{\nu-2}f, \ldots, f, g_1, \ldots, g_{p-1}\}\delta\{z^{\nu-1}f, \ldots, f, g_1, \ldots, g_p\}}{\{z^{\nu-2}f, \ldots, f, g_1, \ldots, g_p\}\delta\{z^{\nu-1}f, \ldots, f, g_1, \ldots, g_{p-1}\}}, \tag{19}$$

where all multigradients are constants and the powers of z merely indicate the functions from which they are formed. In the case of a zero that is distinct in modulus from the others, its reciprocal is the limit with increasing ν of the r's with corresponding subscript. In the case of zeros of equal modulus, the r's enter the coefficients of polynomials in a sequence whose zeros in the limit are these zeros of equal modulus, just as in the qd algorithm [1]. But the simple algorithm for evaluating the r's, which is the qd algorithm itself, is not available in general.

Stewart's own derivation of the qd algorithm as a special case of his is slightly different in that he takes

$$q_j(z) = F^{j-1}g(z).$$

However, by arguments similar to those used to derive the two lemmas, it can be shown that, when this choice is made, then

$$u_p^{(0)}(z) = \{z^{p-2}f, z^{p-3}f, \ldots, f, g, zg, \ldots, z^{p-1}g\},$$

these being elements along a line parallel to and just above the main diagonal in the table of reduced residuals.

REFERENCES

1 Alston S. Householder, *The Numerical Treatment of a Single Nonlinear Equation*, McGraw-Hill Book Co., New York, 1970.
2 Alston S. Householder, The Padé table, the Frobenius identities, and the qd algorithm, *Linear Algebra* 4(1971), 161–174.
3 Alston S. Householder and G. W. Stewart, The numerical factorization of a polynomial, *SIAM Rev.* **13**(1971).
4 G. W. Stewart, On a companion operator for analytic functions, *Numer. Math.* To appear.

Received October 9, 1970